



1/10

A. CTTATCGATAACCGTCGAAACTTGTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCAT
CACAAATTTCACAAATAAAGCATTTCACACTGCATTCTAGTTGTGGTTGTCCAAACTCATCA
+ + + + + ↑ + + + + + + + +
ATGTATCTTATCATGTC CLEAVAGE SITE

B. ————— AAUAAA
+ + + + + GCA

C. ————— GCAaaaaaaaaaaaaaaaaaaaaaa

+ UPSTREAM AND DOWNSTREAM
CLEAVAGE-POLYADENYLATION ELEMENTS

Fig. 1

The diagram illustrates the SV40 genome structure with transcription start sites and regulatory elements. The genome is represented as a horizontal line with various sequence segments and regulatory elements indicated by symbols:

- Transcription Start Sites:** Indicated by upward arrows (↑) above the genome line.
- Enhancer elements:** Indicated by a diamond symbol (◊).
- E2F-motif:** Indicated by a cross symbol (✗).
- Packaging elements:** Indicated by a downward arrow symbol (▼).
- AP3 DNA BS:** Indicated by a triangle symbol (▲).

Key transcription start sites and their corresponding sequences are:

- CATCATCAAT (at position 60)
- TTGTGACCGTG (at position 120)
- GATGTTGCCAA (at position 180)
- GTGTCGCCG (at position 240)
- TAAATTTCGGG (at position 300)
- AGTGAATCT (at position 360)
- GACTTTGACC (at position 420)
- CGGGTCAAAG (at position 480)
- TGAGTTCCCTC (at position 540)
- TCCGACACCG (at position 600)

Regulatory elements include:

- ITR (Indicated by a triangle symbol (▲) at positions 60, 120, and 180)
- AP3 DNA BS (Indicated by a triangle symbol (▲) at position 180)
- TATA Box (Indicated by a triangle symbol (▲) at position 360)
- E1a TATA Box (Indicated by a triangle symbol (▲) at position 420)
- Ar6 (Indicated by a downward arrow symbol (▼) at position 540)
- dl 103-551 (Indicated by a downward arrow symbol (▼) at position 600)
- dl 189-551 (Indicated by a downward arrow symbol (▼) at position 600)
- dl 357-551 (Indicated by a downward arrow symbol (▼) at position 600)



Fig. 2



3/10

1 CATCATCAATAATACCTTATTTGGATTGAAGCCAATATGATAATGAGGGGTGGAGT
+-----ITR-----
61 TTGTGACGTGGCGCGGGCGTGGAACGGGGCGGGTGACGTAGGGCGCGATCAAGCTTAT
+-----ITR-----+-----
121 CGATACCGTCGAAACTGTTATTGCAGCTATAATGGTTACAAATAAAGCAATAGCATC
-----polyA-----
181 ACAAATTTCACAAATAAAGCATTTCAGCTACTGCATTCTAGTTGTGGTTGTCCAAACTC
-----polyA-----
241 ATCAATGTATCTTATCATGTCTGGATCCGCCGCTAGCGATCATCCGGACAAAGCCTGC
-----+-----+-----
301 GCGCGCCCCGCCATTGGCCGTACCGCCCCGGCCGCCCGCCCCATCTGCCCTCG
-----E2F-1 PROMOTER-----
361 CCGCCGGGTCCGGCGCGTTAAAGCCAATAGGAACCGCCGCCGTTGTTCCCACGGCCG
-----E2F-1 PROMOTER-----
421 GGGCAGCCAATTGTGGCGCGCTCGCGGCTCGTGGCTCTTCGCGGAAAAAGGATTG
-----E2f-1 PROMOTER-----
481 GCGCGTAAAAGTGGCCGGACTTGCAGGCAGCGGCCGGGGCGAGCGGGATCGAG
-----E2f-1 PROMOTER-----
541 CCCTCGATGATATCAGATCATGGATCCCGTCGACTGAAAATGAGACATATTATCTGCC
-----+-----+-----
601 ACGGAGGTGTTATTACCGAAGAAATGGCCGCCAGTCTTGACCAGCTGATCGAAGAGG
-----E1a GENE-----
661 TACTGGCTGATAATCTTCCACCTCCTAGCCATTGAAACCACCTACCCCTCACGAACTGT
-----E1a GENE-----
721 ATGATTTAGACGTGACGGCCCCCGAAGATCCAACGAGGAGGCAGTTCGCAGATTTTC
-----E1a GENE-----
781 CCGACTCTGTAATGTTGGCGGTGCAGGAAGGGATTGACTTACTCACTTTCCGCCGGCGC
-----E1a GENE-----
841 CCGGTTCTCCGGAGCCGCCTCACCTTCCGGCAGCCCGAGCAGCCGGAGCAGAGAGCCT
-----E1a GENE-----

Fig. 3A



4/10

901 TGGGTCCGGTTCTATGCCAACCTTGTACCGGAGGTGATCGATCTTACCTGCCACGAGG
-----E1a GENE-----

961 CTGGCTTCCACCCAGTGACGACGAGGATGAAGAGGGTGAGGAGTTGTGTTAGATTATG
-----E1a GENE-----

1021 TGGAGCACCCGGGCACGGTTGCAGGTCTTGTCAATTATCACCGGAGGAATACGGGGGACC
-----E1a GENE-----

1081 CAGATATTATGTGTTCGCTTGCTATATGAGGACCTGTGGCATGTTGTCTACAGTAAGT
-----E1a GENE-----

1141 GAAAATTATGGGCAGTGGGTGATAGAGTGGTGGGTTGGTGTGTAATTTTTTTAAT
-----E1a GENE-----

1201 TTTTACAGTTGTGGTTAAAGAATTTGTATTGTGATTTTTAAAAGGTCTGTGTC
-----E1a GENE-----

1261 TGAACCTGAGCCTGAGCCCGAGCCAGAACCGGAGCCTGCAAGACCTACCCGCCGTCTAA
-----E1a GENE-----

1321 AATGGCGCCTGCTATCCTGAGACGCCGACATCACCTGTGTCTAGAGAATGCAATAGTAG
-----E1a GENE-----

1381 TACGGATAGCTGTGACTCCGGTCCTCTAACACACACCTCCTGAGATAACCCGGTGGTCCC
-----E1a GENE-----

1441 GCTGTCCCCATTAAACCAGTTGCCGTGAGAGTTGGTGGCGCTGCCAGGCTGTGGAATG
-----E1a GENE-----

1501 TATCGAGGACTTGCTTAACGAGCCTGGCAACCTTGGACTTGAGCTGTAAACGCCAG
-----E1a GENE-----

1561 GCCATAAGGTGTAAACCTGTGATTGCGTGTGGTTACGCCTTGTGCTGAATGAGT
-----E1a GENE-----

1621 TGATGTAAGTTAATAAAGGGTGAGATAATGTTAACGGCATGGCGTGTAAATGGGGC
-----+-----

1681 GGGGCTTAAAGGGTATATAATGCGCCGTGGGCTAATCTTGGTTACATCTGACCTCATGGA
-----E1b GENE-----

1741 GGCTGGGAGTGGAAAGATTTCTGCTGTGCGTAACGGCTGGAACAGAGCTCAA
-----E1b GENE-----

1801 CA

Fig. 3B

5/10



33881 AACCTACGCCAGAACGAAAGCCAAAAACCCACAACCTCCTCAAATCGTCACTTCCGT

33941 TTTCCCACGTTACGTACCTCCCATTAAAGAATTCTACAATTCCAACACATACA

34001 AGTTACTCCGCCCTAAAACCCTGGCGAGTCTCACGTAAACGGTCAAAGTCCCCGCGGC
+-- PACKAGING SIGNAL -----

34061 CCTAGACAAATATTACGCGCTATGAGTAACACAAAATTATTCA~~GATTTC~~ACTTCCTCTTA
----- PACKAGING SIGNAL -----

34121 TTCAGTTTCCCGCGAAATGCCAACTTACTCGGTACGCCAAATTACTACAACA
----- PACKAGING SIGNAL -----

34181 TCCGCCTAAAACCGCGCGAAAATTGTCAC~~TCC~~C~~TGT~~TACACCGGCGCACACCAAAAACG
-----+-----

34241 TCACTTTGCCACATCCGTCGCTTACATGTGTTCCGCCACACTGCAACATCACACTCC

34301 GCCACACTACTACGTACCCGCCCGTTCCCACGCCCGGCCACGTACAAACTCCACC
+----- ITR -----

34361 CCCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATATTGATGATG
----- ITR -----+-----

Fig. 3C



6/10

1 CATCATCAATAATACCTATTTGGATTGAAGCCAATATGATAATGAGGGGGTGGAGT
----- ITR -----
61 TTGTGACGTGGCGCGGGCGTGGAACGGGGCGGGTGACGTAGGGCGCGCCGCTAGCGAT
----- ITR -----+-- MCS -----
121 ATCGGATCCCGGTGACTGAAAATGAGACATATTATGCCACGGAGGTGTTATTACCGA
----- E1a -----
181 AGAAATGGCCGCCAGTCTTGACCAGCTGATCGAAGAGGTAATGGCTGATAATCTTCC
----- E1a -----
241 ACCTCCTAGCCATTGAAACCACCTACCCCTCACGAACGTATGATTAGACGTGACGGC
----- E1a -----
301 CCCCCGAAGATCCCAACGAGGAGGCAGCTCGCAGATTTCCCGACTCTGTAATGTTGGC
----- E1a -----
361 GGTGCAGGAAGGGATTGACTTACTCACTTTCCGCCGGCGCCGGTTCTCCGGAGCCGCC
----- E1a -----
421 TCACCTTCCGGCAGCCCGAGCAGCCGGAGCAGAGAGCCTGGTCCGGTTCTATGCC
----- E1a -----
481 AAACCTTGTACCGGAGGTGATCGATCTTACCTGCCACGAGGCTGGCTTCCACCCAGTGA
----- E1a -----
541 CGACGAGGATGAAGAGGGTGAGGAGTTGTGTTAGATTATGTGGAGCACCCGGGACGG
----- E1a -----
601 TTGCAGGTCTTGTCAATTATCACCGGAGGAATACGGGGGACCCAGATATTATGTGTTCGCT
----- E1a -----

Fig. 4



7/10

1 CATCATCAATAATACCTTATTTGGATTGAAGCCAATATGATAATGAGGGGGTGGAGT
+-----ITR-----

61 TTGTGACGTGGCGCGGGCGTGGGAACGGGGCGGGTGACGTAGGGCGCGATCAAGCTTAT
-----ITR-----+-----

121 CGATACCGTCGAAACTGTTATTGCAGCTATAATGGTTACAAATAAGCAATAGCATIC
-----polyA-----

181 ACAAAATTCACAAATAAACGATTTTTCACTGCATTCTAGTTGTGGTTGTCCAAACTC
-----polyA-----

241 ATCAATGTATCTTATCATGTCTGGATCCCGCGCCGCTAGCGATATCGGATCCC GGTCGACT
-----+-----

301 GAAAATGAGACATATTATCTGCCACGGAGGTGTTATTACCGAAGAAATGGCCGCCAGTCT
-----E1a-----

361 TTTGGACCAGCTGATCGAAGAGGTACTGGCTGATAATCTTCCACCTCCTAGCCATTGAA
-----E1a-----

421 ACCACCTACCCTTCACGAACTGTATGATTAGACGTGACGGCCCCCGAAGATCCAACGA
-----E1a-----

481 GGAGGCGGTTTCGCAGATTTCCGACTCTGTAATGTTGGCGGTGCAGGAAGGGATTGA
-----E1a-----

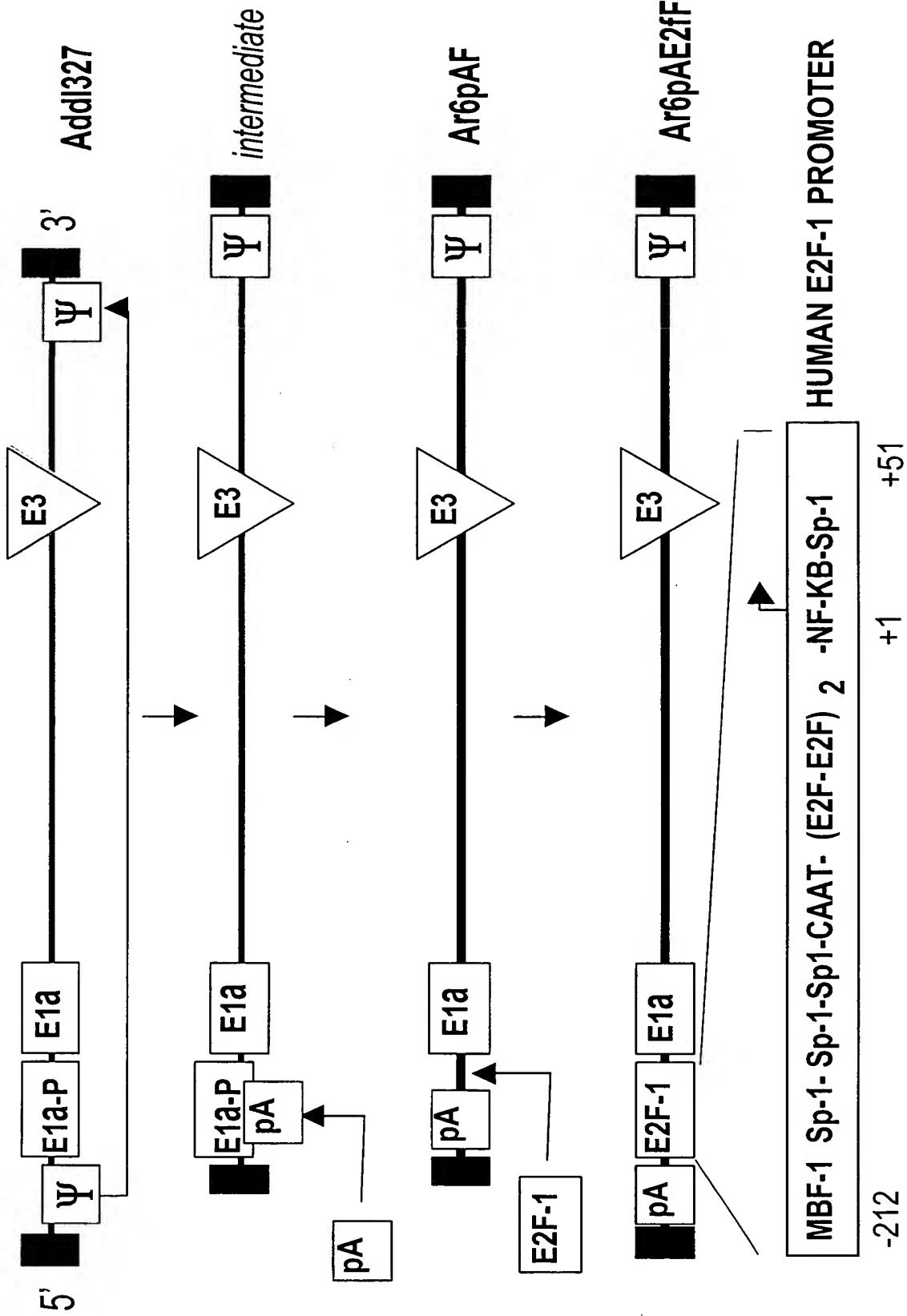
541 CTTACTCACTTCCGCCGGCGCCGGTTCTCCGGAGCCGCCTCACCTTCCGGCAGCC
-----E1a-----

601 CGAGCAGCCGGAGCAGAGAGCCTGGTCCGGTTCTATGCCAACCTTGTACCGGAGGT
-----E1a-----

Fig. 5



8/10





9/10

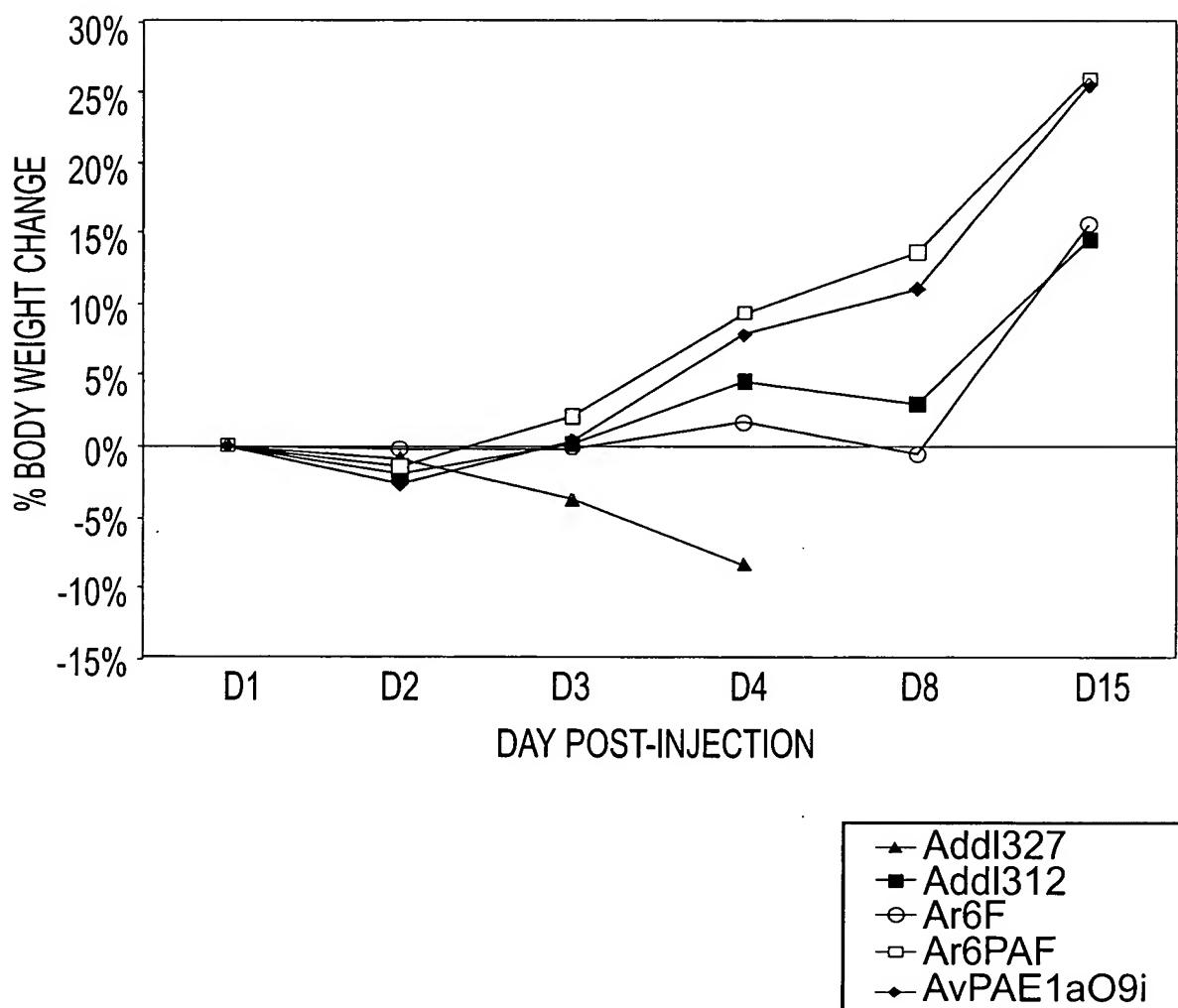


Fig. 7



10/10

Backbones generated:

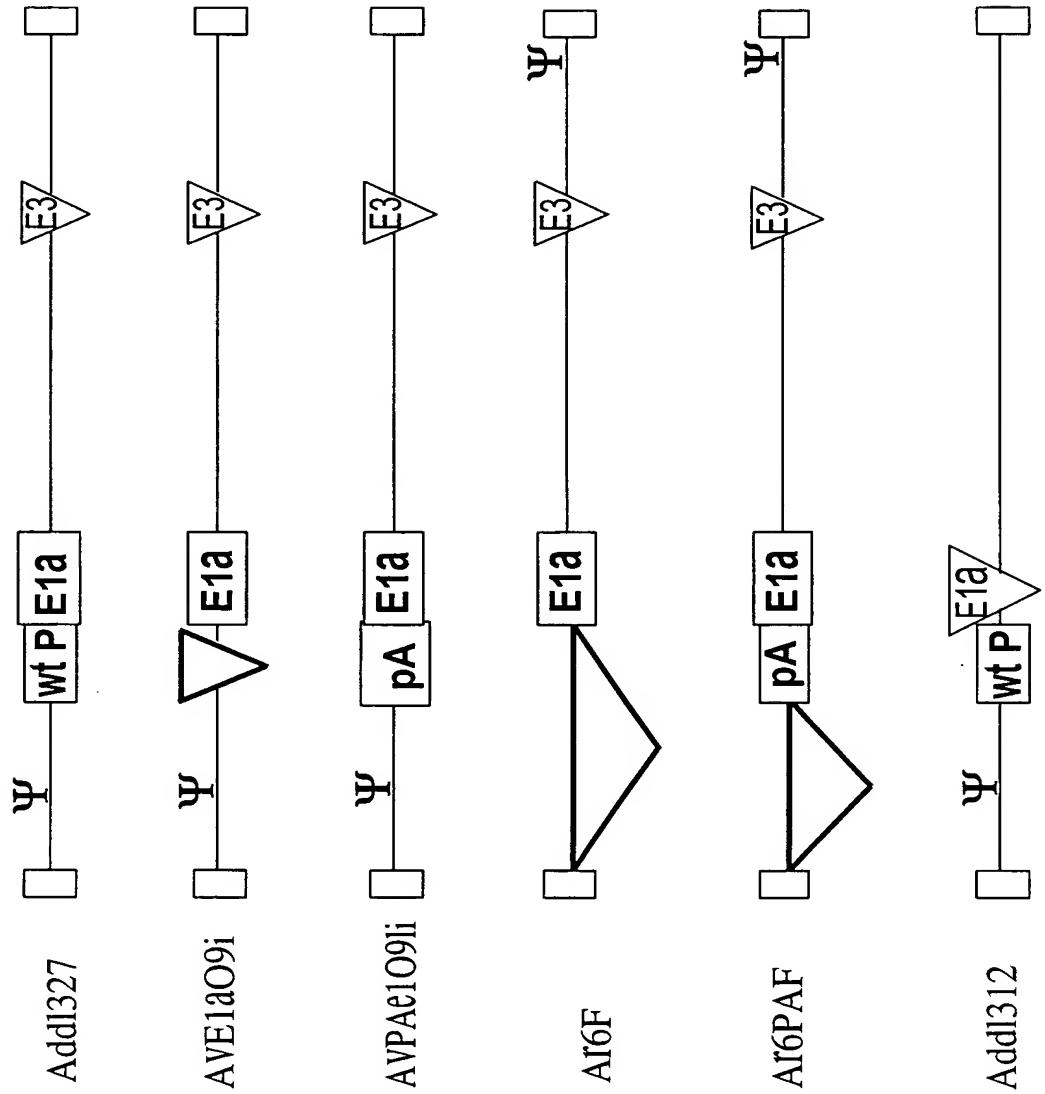


Fig. 8